DOCUMENT RESUME

| | |
|---|---|
| ED 054 204 | TM 000 779 |
| AUTHOR | Pinsky, Paul D. |
| TITLE | Mathematical Models of the Value of Achievement Testing. |
| INSTITUTION | Massachusetts Univ., Amherst. School of Education. |
| SPONS AGENCY | Charles F. Kettering Foundation, Dayton, Ohio. |
| REPORT NO | TM-30 |
| PUB DATE | May 71 |
| NOTE | 27p.; Paper presented at the 39th National Operations Research Society of America Meeting, Dallas, Texas, May 1971 |
| EDRS PRICE | EDRS Price MF-$0.65 HC-$3.29 |
| DESCRIPTORS | Achievement, *Achievement Tests, *Decision Making, Grouping (Instructional Purposes), Individualized Instruction, *Mathematical Models, Measurement, *Statistical Analysis, Teaching Procedures, Testing |
| IDENTIFIERS | *Comprehensive Achievement Monitoring |

ABSTRACT

The mathematical models of this paper were developed as an outgrowth of working with the Comprehensive Achievement Monitoring project (Project CAM) which was conceived as a model and application of sampling procedures such as those used in industrial quality control techniques to educational measurement. This paper explores mathematical modeling techniques to gain insight into the role of achievement testing in the instructional process. All assumptions are listed, three models are presented, and computational examples provided. Each model centers on one of the following: 1) instruction given in a finite amount of time; 2) instructional activities in an individually-paced curriculum; and 3) tests used to group students for instruction. (Author/CK)

Project C omprehensive

A chievement

M onitoring

Technical Memorandum TM-30 May 1971

# MATHEMATICAL MODELS OF THE VALUE OF ACHIEVEMENT TESTING

by

Paul D. Pinsky
Institute for Mathematical Studies in the Social Sciences
Stanford University

(A paper presented at the 39th National Operations Research Society of America Meeting in Dallas, Texas, May, 1971.)

The research and development reported herein was performed pursuant to a grant from the Charles F. Kettering Foundation to the principal investigator, Dwight W. Allen, and under the project director, William P. Gorth, both of the School of Education, The University of Massach setts, Amherst, Massachusetts 01002.



William P. Gorth, Director
Project CAM
School of Education
The University of Massachusetts
Amherst, Massachusetts 01002

# MATHEMATICAL MODELS OF THE VALUE OF ACHIEVEMENT TESTING

Paul D. Pinsky
Institute for Mathematical Studies in the Social Sciences
Stanford University
Stanford, California 94305

A paper presented at the 39th National Operations Research Society of America Meeting in Dallas, Texas, May, 1971

## Contents

## 1.0 Introduction

During the past three years, the author has been working on research and development efforts in educational achievement testing. The Comprehensive Achievement Monitoring project (Gorth [6] and Pinsky [9,10]) that the author is a member of was initiated with the idea of applying sampling procedures such as those used in industrial quality control techniques to educational measurement. Throughout this time, questions have arisen such as how long should an achievement test be?; how frequently should testing be done?; and what is the value of an achievement testing program as compared with its cost? Analogous questions have been asked of the sampling procedures used in industrial quality control techniques. Mathematical models have been developed to gain insight into the role of these sampling procedures in the quality control process (Barlow and Proschan [2]). In his Ph.D dissertation (Pinsky [9]), the author explored the use of these mathematical modeling techniques to gain insight into the role of achievement testing in the instructional process. This paper presents extensions of the work contained in the dissertation.

The next section presents the general assumptions upon which all the models are based. Model 1 is developed in section 3.0; model 2 is developed in section 4.0, followed by computational examples for the model; and model 3 is developed in section 6.0, followed by computational examples for this model. It is the author's intention that these models will be of value in gaining insight into the role of achievement testing in the instructional process. It is not the author's intention that these models will be applied to a real world environment. Much more work is necessary before the latter

situation will be productive.

Some of the assumptions made in this paper are reasonable, and some of the parameters defined are easy to estimate. However, some of the assumptions made and parameters defined should be examined closely. The assumptions about test reliability are straightforward, and the concept of test reliability is easily incorporated in the models. The defining of student parameters should not cause any major problems. The estimation of these parameters is slightly more difficult. The concept of placing a monetary value upon certain instructional activities is a relatively foreign concept in educational institutions. However, this decade should see a considerable effort in the educational community to analyze instructional alternatives in monetary terms. This is evidenced by the Planned Program Budgeting Systems (PPBS) approach that is now getting considerable attention. The assumptions about how the achievement test data is used to make instructional decisions should be examined carefully. The manner in which test data, in conjunction with other data that a teacher has (such as verbal and visual cues in the classroom), is used for decision making is not very well known. It is hoped that these models will stimulate thought about this topic. The final and perhaps most difficult assumption to defend is concerned with the curriculum structure and how a student progresses through this structure. Extensions of the work presented in this paper should concentrate on making more explicit assumptions about the curriculum structure, defining the relevant student parameters, and generating models of learning within this structure.

## 2.0 Assumptions

The following general assumptions apply to all three models. The time variable can be continuous (denoted by the symbol $t$), or discrete (denoted by the symbol $n$). In the assumptions below, time is represented by $t$, but could also be represented by $n$. In each of the models, further assumptions are made.

2.1 Student's state - The student's state at time $t$ is given by $\theta(t)$. $\theta(t)$ can be a scalar or vector quantity, discrete or continuous. Examples of a student's state include what lessons in a textbook he is working on, his grade equivalent in a subject area, his mental ability (IQ) score, or his level on each of the strands in Stanford's CAI mathematics curriculum (Suppes and Morningstar [11]).

2.2 Estimation of $\theta(t_o)$ - Testing is done at time $t_o$. The test results generate an estimate $\hat{\theta}(t_o)$ of the student's true state $\theta(t_o)$. The variance of this estimator is called the reliability of the testing procedure (Lord and Novick [8]).

2.3 Change in the student's state - The student's true state $\theta(t)$ changes over time according to a stochastic learning model. One never observes $\theta(t)$, but must make a projection as to what $\theta(t)$ is given that one knows $\theta(t_o)$. $\theta_p(t)$ is the projected student's state based upon $\theta(t_o)$ and the stochastic learning model. However, one never knows $\theta(t_o)$, but rather must make inferences from $\hat{\theta}(t_o)$. Correspondingly, $\hat{\theta}_p(t)$ is the estimated projected student's state. This variable $\hat{\theta}_p(t)$ is known and is used to make instructional decisions. An example of $\hat{\theta}_p(t)$ is that

a student who has just finished lesson 5 at time $t_o$ $(\hat{\theta}(t_o) = 5)$ will probably have finished lesson 6 by time $t_1$ $(\hat{\theta}_p(t_1) = 6)$ .

2.4 Instructional decisions - Given the value $\hat{\theta}_p(t)$ at time $t$, instructional decisions $D(\hat{\theta}_p(t))$ are made. Thus, if $\hat{\theta}_p(t) = 6$ (meaning that the student has probably just completed lesson 6), then $D(\hat{\theta}_p(t)) = D(6) = 7$ (meaning that the student should work on lesson 7). $D(\cdot)$ can be either a scalar or a vector function.

2.5 Value of the instructional decisions - Given $\theta(t)$, the true value of the student's state; $\hat{\theta}_p(t)$, the estimated projected value of the student's state; and $D(\hat{\theta}_p(t))$, the instruction decisions; $V[D(\theta_p(t))|\theta(t)]$ is the value of such a decision. If the student has just finished lesson 6 $(\theta(t) = 6)$ and is assigned lesson 7 $(D(\hat{\theta}_p(t)) = 7)$, then the value of this instructional decision should be high; while if this student is assigned lesson 12, then the value of the instructional decision should be low.

2.6 Expected value of instruction - It is assumed that testing is done every $T$ units of time. Thus, the models have an infinite horizon. The expected value of the instructional process from time $t_o$ to time $t_o + T$ is given by

$$EV(T) = \int_{t_o}^{t_o+T} \{EV[D(\hat{\theta}_p(t)|\theta(t)]\}dt \qquad \text{for continuous time}$$

and

$$EV(N) = \sum_{n=n_o}^{n_o+N} EV[D(\hat{\theta}_p(n))|\theta(n)] \qquad \text{for discrete time.}$$

The expectation is taken with respect to both the errors of measurement for $\hat{\theta}(t_o)$ and the deviation of $\theta(t)$ from $\theta_p(t)$ .

2.7 Testing frequency - In order to maximize the expected value of the instruction program per unit of time, $T^*$ is chosen which maximizes the expected value of instruction over an interval of $T$ units of time $(EV(T))$ minus the cost of testing $(m)$ divided by $T$ . That is,

$$\frac{EV(T^*) - m}{T^*} = \max_{T > 0} \frac{EV(T) - m}{T}$$

## 3.0 Model 1

The first model is a very simple one. This simplicity is necessary when one considers a finite horizon model; that is, when instruction takes place in a finite amount of time $[0,T]$, and testing is done at times $t_1,\cdots,t_n$, $0 = t_1 \leq \cdots \leq t_n < T$ . Both the finite and infinite horizon formulations of this model are discussed in detail in the author's dissertation (Pinsky [9]).

This section presents a summary of the results for the infinite horizon formulation.

3.1 Student's state - A student is assumed to be in state $c$ (conditioned) or state $uc$ (unconditioned). The terms conditioned and unconditioned are motivated by the terminology of mathematical learning theory (Atkinson, Bowers, and Crothers [1]). The student is assumed to be in state $c$ when the information about his needs is correct, and to be in

state uc when the information about his needs is incorrect. Thus, in state c the student is presented materials such that he learns efficiently, while the student in state uc is presented materials that are not what he really needs (for instance, the materials may be too easy or too hard).

3.2 Estimation of $\theta(t_0)$ - A test is given at time $t_0$ every T units of time. The results of this test are used to decide what the student will study. If the test results are accurate, then $\theta(t_0) = c$ . If the test results are inaccurate, then $\theta(t_0) = uc$ . After each test, the probability that a student will be in state c is equal to r, i.e., $P\{\theta(t_0) = c\} = r$ . r is referred to as the reliability of the testing procedure. One never knows what state the student is in. In this model, $\hat{\theta}(t_0) = c$ .

3.3 Change in the student's state - As the length of time increases since the last test, the information about the student's needs becomes less accurate. Correspondingly, the materials presented to the student are less in accord with his needs. A student who is in state c at time $t_0$ might switch to state uc at time $t_0 + t$ . A student in state uc remains in state uc until at least the time when the next test is administered. This transition of student states is summarized below.

$$P\{\theta(t) = c|\theta(t_0) = c\} = 1 - F(t - t_0) \qquad P\{\theta(t) = c|\theta(t_0) = uc\} = 0$$

$$P\{\theta(t) = uc|\theta(t_0) = c\} = F(t - t_0) \qquad P\{\theta(t) = uc|\theta(t_0) = uc\} = 1$$

$F(\cdot)$ is referred to as the failure distribution analogous to similar terminology used in models of the failure of mechanical components in industrial situations (Barlow and Proschan [2]). Later in this section, two forms of the failure distribution will be used:

(a) uniform $F(t) = \min[\frac{t}{D},1]$, $t \geq 0$

(b) exponential $F(t) = 1 - e^{-\lambda t}$, $t \geq 0$ .

One never knows if the student's true state changes. In this model, $\hat{\theta}_p(t) = c$ all the time.

3.4 Instructional decisions - Since $\hat{\theta}_p(t) = c$, the instructional decisions are always the same - $D(\hat{\theta}_p(t)) = D(c) \equiv D_c$ .

3.5 Value of the instructional decisions- The value of the instructional decisions is given by

$$V[D(\hat{\theta}_p(t))|\theta(t)] = V[D_c|\theta(t)] = \begin{cases} 1 & \text{if } \theta(t) = c \\ \alpha & \text{if } \theta(t) = uc,\ 0 \leq \alpha < 1 \end{cases}$$

Thus, a student in state c learns at a rate of 1 per unit time, while a student in state uc learns at a rate of $\alpha$ per unit time.

3.6 Expected value of instruction - The value of the instructional program has two sources of variance. The first is the probability of initially being placed in state c, and the second is the probability of transferring from state c to state uc at time t . With $t_o = 0$, one obtains

$$EV(t) = \int_0^T [P\{0(t) = c\} \cdot 1 + P\{0(t) = uc\} \cdot \alpha]dt$$

$$P\{0(t) = c\} = r \cdot (1 - F(t))$$

$$P\{0(t) = uc\} = (1 - r) + rF(t)$$

thus

$$EV(t) = [\alpha(1 - r) + r]T - r(1 - \alpha) \int_0^T F(t)dt \quad .$$

3.7 Testing frequency -

(a) Uniform failure distribution

$$EV(T) = [\alpha(1 - r) + r]T - \frac{r(1 - \alpha)T^2}{2D} \qquad \text{for} \quad T \leq D$$

$$\frac{EV(T) - m}{T} = \alpha(1 - r) + r - \frac{r(1 - \alpha)T}{2D} - \frac{m}{T}$$

Taking the derivative of the above with respect to $T$, setting it equal to zero, and solving for $T$ yields

$$T^* = \left(\frac{2mD}{r(1 - \alpha)}\right)^{\frac{1}{2}}$$

This solution only holds for $T^* \leq D$ .

(b) Exponential failure distribution

$$EV(T) = [\alpha(1 - r) + r]T - r(1 - \alpha)[T + \frac{e^{-\lambda T} - 1}{\lambda}]$$

$$\frac{EV(T) - m}{T} = \alpha(1 - r) + r - r(1 - \alpha) - r(1 - \alpha)\frac{(e^{-\lambda T} - 1)}{\lambda T} - \frac{m}{T}$$

Taking the derivative of the above with respect to $T$ and setting it equal to zero, yields the following equation for $T^*$

$$(1 + \lambda T^*)e^{-\lambda T^*} = 1 - \frac{\lambda m}{(1 - \alpha)r} \quad .$$

The author has been unable to find any closed form solution of $T^*$ from this equation.

## 4.0 Model 2

The second model is concerned with tests that are used to direct individual students' instructional activities in an individually-paced curriculum. Examples of this type of instructional program include Individually Prescribed Instruction- IPI (Cooley and Glaser [3]), Planned Learning in Accord with Needs - PLAN (Flanagan [5]), and Stanford's CAI strands curriculum (Suppes and Morningstar [11]).

4.1 Student's state - In this model the student's state is a one-dimensional variable that is related to his achievement level in the subject area. In the computational example presented in the next section, the curriculum content ranges from kindergarten to the sixth grade, and a

student's state can range from 000 (the entry level in the kindergarten) to 699 (the exit level from the sixth grade). Thus, the state variable is approximately a grade equivalent achievement level for a student.

4.2 Estimation of $\theta(t_o)$ - A test is given to a student at time $t_o$ every $T$ units of time. $\hat{\theta}(t_o)$ is the estimate of $\theta(t_o)$ that is generated by the test data. It is assumed that $\hat{\theta}(t_o) \sim N(\theta(t_o),\sigma_r^2)$. $\sigma_r^2$ is referred to as the reliability of the test.

4.3 Change in the student's state - $\theta(t)$ is assumed to change over time according to the following linear learning model.

$$\theta(t) = \theta(t_o) + \alpha(t - t_o) + \varepsilon(t - t_o) \qquad t \geq t_o$$

where $\varepsilon(t - t_o) \sim N(0,\sigma_\varepsilon^2(t - t_o))$ .
$\theta_p(t)$ = the projected student's state = $E(\theta(t)) = \theta(t_o) + \alpha(t - t_o)$ .
Thus, $\theta(t) \sim N(\theta_p(t), \sigma_\varepsilon^2(t - t_o))$. $\sigma_\varepsilon^2$ is an indicator of how accurately one can predict a student's achievement level over time in the curriculum structure. $\hat{\theta}_p(t)$ = the estimated projected student's state = $\hat{\theta}(t_o) + \alpha(t - t_o) \sim N(\theta_p(t),\sigma_r^2)$ . Define a new variable $\theta_d(t) = \theta(t) - \hat{\theta}_p(t) \sim N(0,\sigma_r^2 + \sigma_\varepsilon^2(t - t_o))$ . This variable is the difference between the student's true state and the estimated projected student's state.

4.4 Instructional decisions - Given knowledge that the student's true state $\theta(t) = X$, one would want to provide the student with instructional materials related to the curriculum content slightly more difficult than $X$ . An instructional decision of $X$ means providing materials rela-

ted to the curriculum content slightly more difficult than X. Thus $D(\hat{\theta}_p(t)) = \hat{\theta}_p(t)$.

4.5 Value of the instructional decisions - The value of instructional decision X for a student in state Y depends upon the closeness of X and Y. The value of presenting a lesson on curriculum content 425-430 to a student in state 425 is high, while the value of presenting this lesson to a student of state 300 or state 600 is quite low. The value function chosen for this model is

$$V[D(\hat{\theta}_p(t))|\theta(t)] = ae^{-b(\theta(t) - \hat{\theta}_p(t))^2} .$$

4.6 Expected value of instruction - Assume $t_0 = 0$ for the remainder of this section. The expectation of the value of instruction is taken with respect to two variables, $\theta(t)$ and $\hat{\theta}(t_0)$

$$EV[D(\hat{\theta}_p(t))|\theta(t)] = E_{\theta(t)}\left\{E_{\hat{\theta}(t)}\left[ae^{-b(\theta(t) - \hat{\theta}_p(t))^2}\right]\right\} .$$

Using the variable transformation $\theta_d(t) = \theta(t) - \hat{\theta}_p(t)$ and using the properties of the normal distribution

$$EV[\cdot|\cdot] = \int_{-\infty}^{\infty} ae^{-b\theta_d(t)} f(\theta_d(t))\, d\theta_{d(t)} .$$

Since $\upsilon_d(t) \sim N(0,\sigma_r^2 + \sigma_\varepsilon^2 t)$, the above equals

$$= a[2b(\sigma_r^2 + \sigma_\varepsilon^2 t) + 1]^{-1/2} \quad .$$

Then

$$EV(T) = \int_0^T a[2b(\sigma_r^2 + \sigma_\varepsilon^2 t) + 1]^{1/2} dt \quad .$$

Using the variable transformation $Y = 2b(\sigma_r^2 + \sigma_\varepsilon^2 t) + 1$, one obtains

$$EV(T) = \frac{a}{b\sigma_\varepsilon^2}\left\{[2b(\sigma_r^2 + \sigma_\varepsilon^2 T]^{1/2} - [2b\sigma_r^2 + 1]^{1/2}\right\} \quad .$$

4.7 Testing frequency - The optimal testing frequency $T^*$ in this model is that value of $T$ which maximizes

$$(4.7.1) \qquad \frac{EV(T) - m}{T} = \frac{\frac{a}{b\sigma_\varepsilon^2}\left\{[2b(\sigma_r^2 + \sigma_\varepsilon^2 T) + 1]^{1/2} - [2b\sigma_r^2 + 1]^{1/2}\right\} - m}{T} \quad .$$

The author could find no closed form solution for $T^*$ . The next section presents a computational example of the above function.

## 5.0 Computational Example for Model 2

A computer program was written to find $T^*$, the optimal testing frequency for model 2, as a function of the various parameters of the model.

Upon examining equation (4.7.1), one notices that $T^*$ is a function of only three variables, $m/a$, $b\sigma_\varepsilon^2$, and $b\sigma_r^2$. The following thoughts went into the selection of the parameter values used in the examples.

(a) The student state and curriculum content are numbered from 000 to 699. 000 is the entry level for kindergarten, while 699 is the exit level from the sixth grade.

(b) Time is in school days. 180 school days to the year times 7 school years equals 1260 school days from kindergarten through the sixth grade.

(c) If the average student enters level 0 and leaves level 699 over a seven year period, then $\alpha = .55$.

(d) \$100-\$150 per student per subject area per year is approximately what most school districts spend. Thus a maximum value of \$1 per day for instruction in a subject area was chosen. $a = 1$

(e) $\sigma_r^2 = 1$ implies an excellent test while $\sigma_r^2 = 100$ implies a very poor test.

(f) $\sigma_\varepsilon^2 = .02$ implies a curriculum and student population that allows for very accurate prediction of student achievement, while $\sigma_\varepsilon^2 = 2.0$ implies a curriculum and student population for which accurate prediction is not possible.

(g) A value of $b = .05$ was chosen.

(h) The cost per test per student ($m$) ranges from \$.10 to \$.90.

The results of the computer analysis are shown in Table 1. For each of the 5 values of $b\sigma_\varepsilon^2$, $b\sigma_r^2$, and $m/a$, the computer calculated equation (4.7.1) for $T = 1$ to 100 in steps of 1 and found the optimal testing frequency $T^*$. The number in the left of the brackets is $T^*$,

while the number in the right of the brackets is the value (in cents per day) of the instructional program with a testing frequency of $T^*$ . A $*$ in the brackets means that $T^*$ was not found in the calculations, i.e., $T^* \geq 100$ . Table 1 allows one to do a sensitivity analysis of the parameters in model 2. Note that the value of the instructional program seems to be more sensitive to the reliability of the testing procedures $(b\sigma_r^2)$ than to the cost of testing $(m)$ or the predictive power of the student's achievement level $(b\sigma_\varepsilon^2)$ . Figure 1 contains a graph of $\frac{EV(T) - m}{T}$ as a function of $T$ for $a = 1$, $b = .05$, $b\sigma_\varepsilon^2 = .03$, $b\sigma_r^2 = .5$ and $m = .70$ .

## 6.0 Model 3

The third model is concerned with tests that are used to group students together for instructional activities. For each unit of curriculum content, there may be several different instructional activities. In many instances, these activities vary in their difficulty level. Examples of this type of instructional program include tracking systems that place each student in a certain difficulty level of instructional activities and periodically change the student's level when his achievement results change, and Stanford's CAI block structure curriculum (Jerman and Suppes [7]). The time variable in this model is discrete and is labeled as $n = 1,2,3,\cdots$ This time variable corresponds to school days.

6.1 Student's state - During each time period (or day), the student is assumed to be in one of $S$ states, i.e., $\theta(n) \varepsilon \{1,\cdots,S\}$ . This state

variable corresponds to the instructional activity that is best suited for the student's needs. There are $S$ instructional activities available in the model.

6.2 Estimation of $\theta(n_o)$ - A test is given to each student at the end of time period $n_o$ every $N$ units of time. The test results are used to place the student in the instructional activity that appears best suited for his needs. Thus $\hat{\theta}(n_o) \in \{1,\cdots,S\}$ . Given that a student is actually in state $\theta(n_o) = j$, the test will place him in state $\hat{\theta}(n_o) = k$ with probability $r_{jk}$, $\sum_{k=1}^{S} r_{jk} = 1$, $j = 1,\ldots,S$ . $(r_{jk})$ is referred to as the reliability matrix of the testing procedure.

6.3 Change in the student's state - A student in state $j$ at time $n$ transfers to state $\ell$ at time $n+1$ with probability $p_{j\ell}$, $\sum_{\ell=1}^{S} p_{j\ell} = 1$, $j = 1,\ldots,S$ . $(p_{j\ell})$ is referred to as the student transition matrix. The states are labeled such that $p_{jj} \geq p_{j\ell}$, $\ell = 1,\ldots,S$ . A student who is assigned instructional activity $k$ at time $n_o$, remains in activity $k$ until the next test. Since testing is done every $N$ units of time, $\hat{\theta}_p(n_1 + n_o) = \hat{\theta}_p(n_o) = \hat{\theta}(n_o)$, $n_1 = 0,\ldots,N-1$ .

6.4 Instructional decisions - A student with parameter $\hat{\theta}_p(n) = k$ is assigned instructional activity $k$; that is, $D(\hat{\theta}_p(n)) = \hat{\theta}_p(n)$ . Testing is done to determine what activity is best for each student, and the student is kept in that activity until another test is given.

6.5 Value of the instructional decision - A student in state $\ell$ $(\theta(n) = \ell)$, and in instructional activity $k$, $(\hat{\theta}_p(n) = k)$ has a value $V_{\ell k}$;

that is,

$$V[D(\hat{o}_p(n)) = k \mid o(n) = \ell] = V_{\ell k} \quad .$$

It is assumed that $V_{\ell\ell} \geq V_{\ell k}$, $k = 1,\ldots,S$ .

6.6 Expected value of instruction - In order to calculate the expected value of instruction, one must know the percentage of students who are in each of the $S$ states. Let $\pi_j$ be the percentage of students in state $j$ . Since the model being considered is an infinite horizon one, the $\pi_j$ should be the steady state probabilities of the student transition matrix $(P_{j\ell})$ . (This model contains a finite state Markov chain. The properties of a finite state Markov chain that are used in this paper came from Chapter XVI of Fellar [4].) After a test is given and the student assigned to instructional activities, the value per unit time is given by

$$\sum_{j=1}^{S} \pi_j \sum_{k=1}^{S} r_{jk} V_{jk} \quad .$$

This expression results from a student in state $j$ (which has probability $\pi_j$ ) being assigned to activity $k$ (which has probability $r_{jk}$ ) which has a value of $V_{jk}$ . Between time periods $n_o + 1$ and $n_o + 2$, the student's state changes from state $j$ to state $\ell$ with probability $p_{j\ell}$ . The student, however, remains in activity $k$ . Thus the value per unit time at time $n_o + 2$ is

$$\sum_{j=1}^{S} \pi_j \sum_{k=1}^{S} r_{jk} \sum_{\ell=1}^{S} p_{j\ell} V_{\ell k}$$

In general, the value of the instructional program per unit time $n$ units of time after testing is given by

$$(6.6.1) \qquad \sum_{j=1}^{S} \pi_j \sum_{j=1}^{S} r_{jk} \sum_{\ell=1}^{S} p_{j\ell}^{n-1} V_{\ell k}$$

where $p_{j\ell}^{n-1}$ is the probability of a student transferring from state $j$ to state $\ell$ in $n-1$ time intervals. Thus, the expected value of an instruction program when testing is done every $N$ time units is

$$EV(N) = \sum_{n=n_o+1}^{n_o+N} \left\{ \sum_{j=1}^{S} \pi_j \sum_{k=1}^{S} r_{jk} \sum_{\ell=1}^{S} p_{j\ell}^{n-1} V_{\ell k} \right\}$$

6.7 Testing frequency - The optimal testing frequency $N^*$ under the assumptions of model 3 is that $N$ which maximizes

$$(6.7.1) \qquad \frac{EV(N) - m}{N} = \frac{\sum_{n=n_o+1}^{n_o+N} \left\{ \sum_{j=1}^{S} \pi_j \sum_{k=1}^{S} r_{jk} \sum_{\ell=1}^{S} p_{j\ell}^{n-1} V_{\ell k} \right\} - m}{N}$$

The next section presents a computational example of the above function for the case $S = 2$ .

## 7.0 Computational Example for Model 3

Because of the complexity of equation (6.7.1) and because of the number of parameters involved, $S = 2$ was chosen to generate some data for

model 3. With $S = 2$, the following reduction in the number of parameters occurs.

$$r_{11} = 1 - r_1, \; r_{12} = r_1, \; r_{21} = r_2, \; r_{22} = 1 - r_2 \; .$$

Thus, $r_j$ is the probability of assigning a student in state $j$ to activity $k$, $k \neq j$ .

$$p_{11} = 1 - \alpha_1, \; p_{12} = \alpha_1, \; p_{21} = \alpha_2, \; p_{22} = 1 - \alpha_2 \; .$$

Thus, $\alpha_j$ is the probability of a student in state $j$ transferring to state $k$, $k \neq j$, in one time interval.

$$\pi_1 = \frac{\alpha_2}{\alpha_1 + \alpha_2} , \qquad \pi_2 = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

$\pi_1$ and $\pi_2$ are the steady state probabilities of a student being in each state.

$$p_{11}^{n-1} = \frac{\alpha_2}{\alpha_1 + \alpha_2} + \frac{\alpha_1(1-\alpha_1-\alpha_2)^n}{\alpha_1 + \alpha_2} ; \qquad p_{12}^{n-1} = \frac{\alpha_1}{\alpha_1 + \alpha_2} - \frac{\alpha_1(1-\alpha_1-\alpha_2)^n}{\alpha_1 + \alpha_2}$$

$$p_{21}^{n-1} = \frac{\alpha_2}{\alpha_1 + \alpha_2} - \frac{\alpha_2(1-\alpha_1-\alpha_2)^n}{\alpha_1 + \alpha_2} ; \qquad p_{22}^{n-1} = \frac{\alpha_1}{\alpha_1 + \alpha_2} + \frac{\alpha_2(1-\alpha_1-\alpha_2)^n}{\alpha_1 + \alpha_2}$$

The derivation of the above expressions can be found in Chapter XVI of Fellar [4]. Inserting the above parameters into equation (6.6.1) and reducing yields,

$$K_1 + (\alpha_1 + \alpha_2)K_2(1 - \alpha_1 - \alpha_2)^{n-1}$$

where

$$(\alpha_1 + \alpha_2)^2 K_1 = V_{11}[\alpha_2^2(1 - r_1) + \alpha_1\alpha_2 r_2] + V_{12}[\alpha_2^2 r_1 + \alpha_1\alpha_2(1 - r_2)]$$

$$+ V_{21}[\alpha_1\alpha_2(1 - r_1) + \alpha_1^2 r_2] + V_{22}[\alpha_1\alpha_2 r_1 + \alpha_2^2(1 - r_2)]$$

and

$$(\alpha_1 + \alpha_2)^3 K_2 = \alpha_1\alpha_2(1 - r_1 - r_2)[V_{11} + V_{22} - V_{12} - V_{21}]$$

then

$$EV(N) = K_1 + K_2[1 - (1 - \alpha_1 - \alpha_2)^N] \quad .$$

Finally, equation (6.7.1) becomes (with $n_o = 0$)

$$\frac{EV(N) - m}{N} = \frac{K_1 + K_2[1 - (1 - \alpha_1 - \alpha_2)^N] - m}{N}$$

Note that $N^*$, the optimal value of $N$, depends only upon $\alpha_1$, $\alpha_2$, $r_1 + r_2$, and $m/[V_{11} + V_{22} - V_{12} - V_{21}]$, while $\frac{EV(N^*) - m}{N^*}$ depends upon the values of $\alpha_1$, $\alpha_2$, $r_1$, $r_2$, $V_{11}$, $V_{12}$, $V_{21}$, $V_{22}$, and $m$ .

The following thoughts went into the selection of the values used in the examples.

(a) Time is in terms of school days.

(b) \$100-\$150 per student per subject area per year is approximately

what most school districts spend. $V_{11} = V_{22} = 1$ and $V_{12} = V_{21} = \frac{1}{2}$ were chosen. This means that $V_{11} + V_{22} - V_{12} - V_{21} = 1$ .

(c) $r_1$ and $r_2$ range from .01 for an excellent testing program to .33 for a very poor testing program.

(d) $\alpha_1$ and $\alpha_2$ range from .01 for a very stable student population and curriculum structure to .25 for a very unstable student population and curriculum structure.

(e) The cost per test per student (m) ranges from \$.10 to \$.90.

The results of the computer analysis are shown in Table 2. For each of the values of the parameters the computer calculated $\frac{EV(N) - m}{N}$ for $N = 1$ to 100 in steps of 1 and found the optimal testing frequency $N^*$ . The number in the left of the brackets is $N^*$ . The number in the right of each bracket is the value (in cents per day) of the instructional program with a testing frequency of $N^*$ . A * in the brackets means that $N^*$ was not found in the calculations, i.e., $N^* \geq 100$ . Figure 2 contains a graph of $\frac{EV(N) - m}{N}$ as a function of N for $m = .3$, $\alpha_1 = \alpha_2 = .05$, and $r_1 + r_2 = .25$ .

## References


1. Atkinson, R. C., Bower, G. H., and Crothers, E. J., An Introduction to Mathematical Learning Theory, John Wiley and Sons, New York, 1965.

2. Barlow, R. E., and Proschan, F., Mathematical Theory of Reliability, John Wiley and Sons, 1965.

3. Cooley, William W., and Glaser, Robert, "An Information and Management System for Individually Prescribed Instruction", Learning Research and Development Center, University of Pittsburgh, December, 1968.

4. Fellar, William, An Introduction to Probability Theory and Its Applications, John Wiley and Sons, New York, 1957.

5. Flanagan, John C., "Functional Education for the Seventies", Phi Delta Kappan, September, 1967, 27-33.

6. Gorth, W. P., Schriber, P. E., and O'Reilly, R. P., Training Manual for Comprehensive Achievement Monitoring, The New York State Education Department, Division of Research, Albany, N. Y., June, 1970.

7. Jerman, Max, and Suppes, Patrick, "A Workshop on Computer-Assisted Instruction in Elementary Mathematics", The Arithmetic Teacher, March, 1969, 193-197.

8. Lord, F. M., and Novick, M. R., Statistical Theories of Mental Test Scores, Addison-Wesley, Boston, 1968.

9. Pinsky, Paul D., "Mathematical Models Concerning the Measurement of Educational Achievement Levels", Ph.D. dissertation, Department of Operations Research, Stanford University, May, 1971.

10. Pinsky, Paul D., "Feedback Information in the Comprehensive Monitoring of Educational Achievement", A paper presented at the 11th American Meeting of the Institute of Management Sciences, Los Angeles, October, 1970.

11. Suppes, Patrick, and Morningstar, Mona, "Technical Inovations : Computer-Assisted Instruction and Compensatory Education", in F. Korten, S. Cook, and J. Lacey (Eds.), Psychology and the Problems of Society, American Psychological Association, Inc., Washington, D. C., 1970.

| $b\sigma_\varepsilon^2$ | .001 | .003 | .010 | .030 | .10 |
|---|---|---|---|---|---|
| $b\sigma_r^2 = .05$ | | | | | |
| m = .1a | [15, 94] | [9, 93] | [5, 91] | [3, 88] | [2, 83] |
| m = .3a | [27, 93] | [16, 91] | [9, 88] | [6, 84] | [3, 75] |
| m = .5a | [36, 92] | [21, 90] | [12, 86] | [7, 81] | [5, 70] |
| m = .7a | [43, 92] | [25, 90] | [14, 85] | [9, 78] | [6, 66] |
| m = .9a | [48, 91] | [28, 89] | [16, 84] | [11, 76] | [7, 63] |
| $b\sigma_r^2 = .15$ | | | | | |
| m = .1a | [18, 87] | [10, 86] | [6, 84] | [3, 82] | [2, 77] |
| m = .3a | [30, 86] | [18, 84] | [10, 82] | [6, 77] | [4, 70] |
| m = .5a | [40, 85] | [23, 83] | [13, 80] | [8, 75] | [5, 65] |
| m = .7a | [48, 85] | [28, 83] | [16, 79] | [10, 72] | [7, 62] |
| m = .9a | [54, 84] | [32, 82] | [19, 77] | [12, 71] | [8, 59] |
| $b\sigma_r^2 = .50$ | | | | | |
| m = .1a | [25, 70] | [14, 69] | [8, 68] | [5, 66] | [3, 63] |
| m = .3a | [42, 69] | [25, 68] | [14, 66] | [8, 63] | [5, 58] |
| m = .5a | [56, 69] | [32, 68] | [18, 65] | [11, 61] | [7, 54] |
| m = .7a | [66, 69] | [38, 67] | [22, 64] | [14, 60] | [9, 52] |
| m = .9a | [75, 68] | [44, 66] | [25, 63] | [16, 58] | [10, 50] |
| $b\sigma_r^2 = 1.50$ | | | | | |
| m = .1a | [39, 50] | [24, 49] | [13, 48] | [8, 47] | [4, 45] |
| m = .3a | [70, 49] | [42, 49] | [23, 47] | [14, 45] | [8, 42] |
| m = .5a | [91, 49] | [53, 48] | [30, 47] | [18, 44] | [11, 40] |
| m = .7a | * | [64, 48] | [36, 46] | [22, 43] | [14, 38] |
| m = .9a | * | [72, 47] | [42, 45] | [26, 42] | [16, 37] |
| $b\sigma_r^2 = 5.00$ | | | | | |
| m = .1a | [86, 30] | [51, 30] | [28, 29] | [16, 29] | [9, 28] |
| m = .3a | * | [88, 29] | [49, 29] | [29, 28] | [17, 26] |
| m = .5a | * | * | [64, 29] | [38, 27] | [23, 25] |
| m = .7a | * | * | [76, 28] | [46, 27] | [28, 25] |
| m = .9a | * | * | [87, 28] | [53, 27] | [33, 24] |

[T*, the optimal testing frequency, The optimal value of the instructional program in cents per day]

Computational Example for Model 2

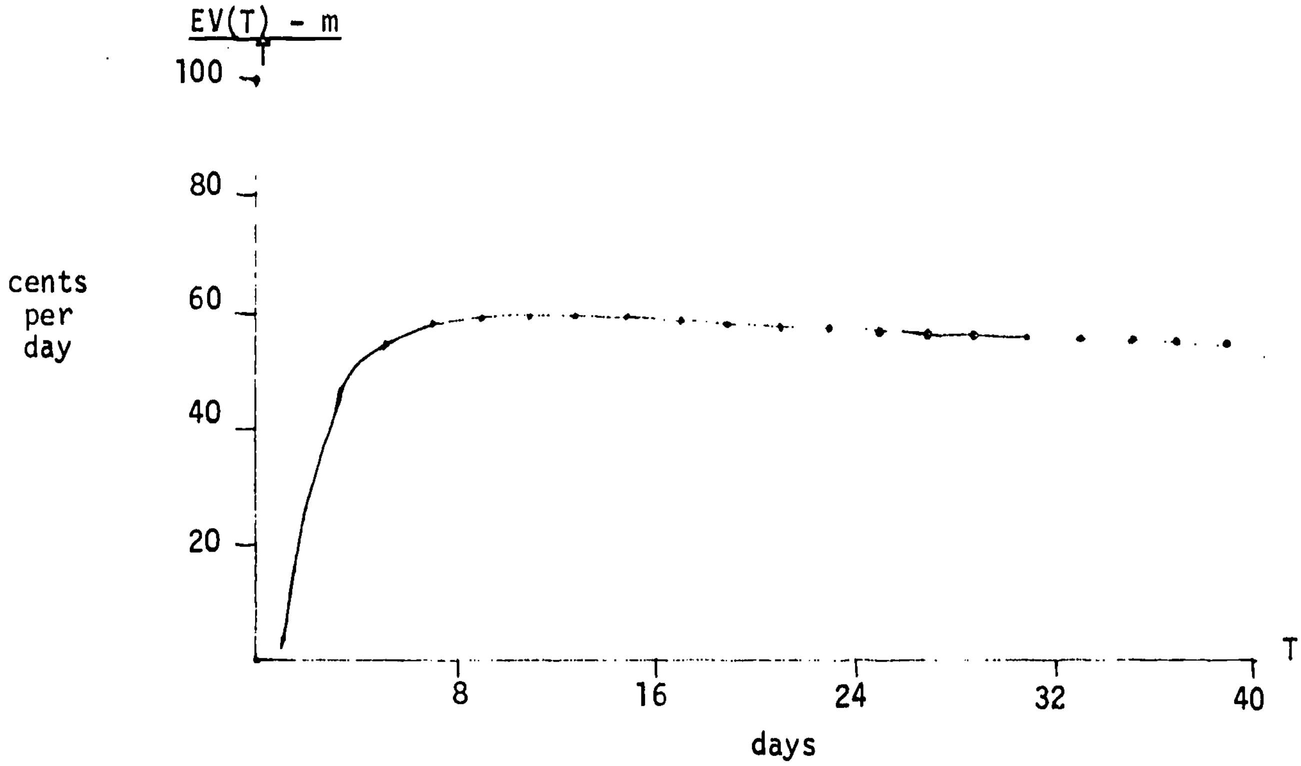
Table 1

| $(\alpha_1, \alpha_2) =$ | (.01,.01) | (.01,.05) | (.01,.25) | (.05,.05) | (.05,.25) | (.25,.25) |
|---|---|---|---|---|---|---|
| **$r_1 + r_2 = .02$** | | | | | | |
| m = .1 | [7, 97] | [5, 96] | [8, 96] | [3, 94] | [3, 92] | [1, 89] |
| m = .3 | [12, 94] | [10, 93] | * | [6, 89] | [6, 87] | [3, 79] |
| m = .5 | [16, 93] | [14, 92] | * | [8, 86] | * | * |
| m = .7 | [19, 92] | [18, 91] | * | [10, 84] | * | * |
| m = .9 | [22, 91] | [22, 90] | * | [12, 82] | * | * |
| **$r_1 + r_2 = .10$** | | | | | | |
| m = .1 | [7, 95] | [6, 93] | [10, 93] | [3, 92] | [3, 91] | [1, 87] |
| m = .3 | [12, 93] | [11, 91] | * | [6, 87] | [7, 86] | [3, 78] |
| m = .5 | [16, 91] | [15, 89] | * | [8, 85] | * | * |
| m = .7 | [20, 90] | [19, 88] | * | [11, 83] | * | * |
| m = .9 | [23, 89] | [24, 87] | * | [13, 81] | * | * |
| **$r_1 + r_2 = .25$** | | | | | | |
| m = .1 | [8, 91] | [6, 90] | [15, 90] | [4, 89] | [3, 88] | [2, 84] |
| m = .3 | [14, 89] | [12, 88] | * | [7, 85] | [10, 84] | [4, 76] |
| m = .5 | [18, 88] | [17, 87] | * | [10, 82] | * | * |
| m = .7 | [22, 87] | [22, 86] | * | [12, 80] | * | * |
| m = .9 | [26, 86] | [28, 85] | * | [15, 79] | * | * |
| **$r_1 + r_2 = .67$** | | | | | | |
| m = .1 | [12, 82] | [10, 81] | * | [6, 80] | [6, 79] | [3, 76] |
| m = .3 | [22, 81] | [22, 79] | * | [12, 77] | * | * |
| m = .5 | [30, 80] | [36, 79] | * | [19, 76] | * | * |
| m = .7 | [37, 79] | [66, 78] | * | [32, 75] | * | * |
| m = .9 | [43, 78] | * | * | [*, 75] | * | * |

[N*, the optimal testing frequency, The optimal value of the instructional program in cents per day]

Computational Example for Model 3

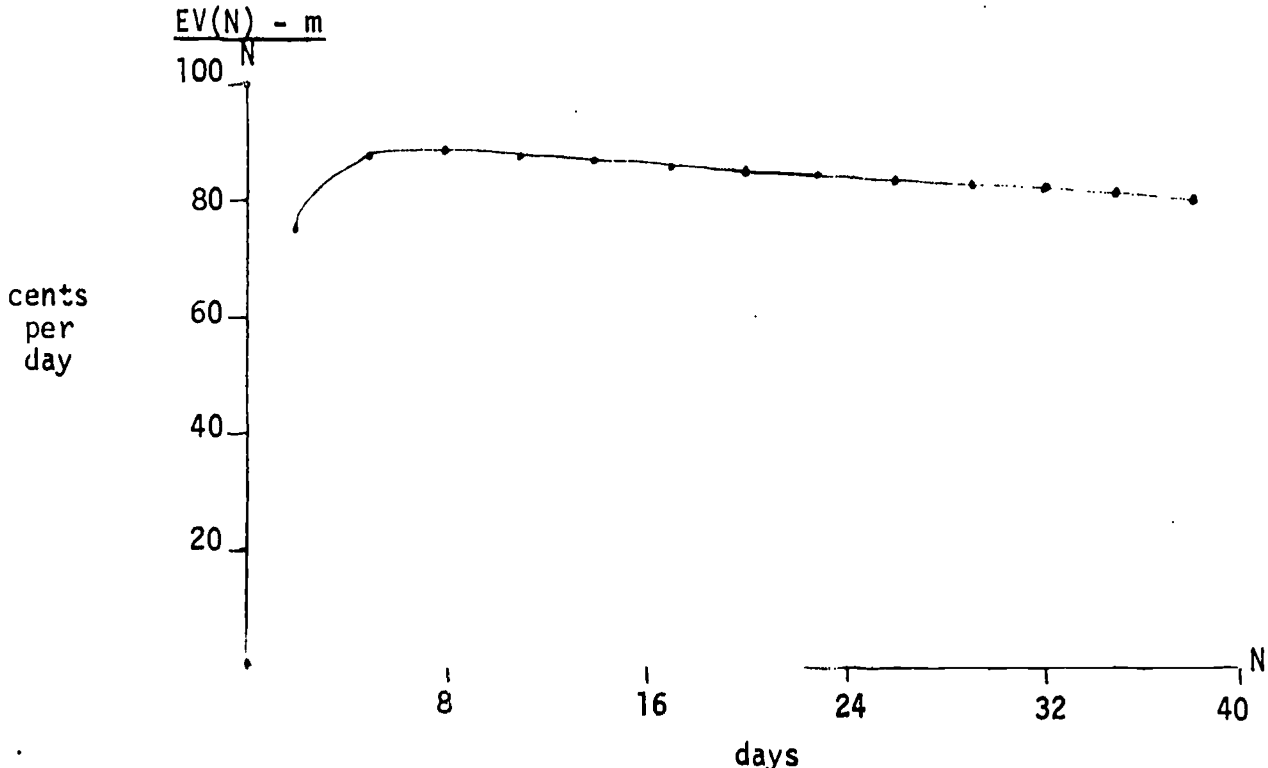
Table 2




Value of the instruction program as a function of testing frequency

Model 2

Figure 1




Value of the instruction program as a function of testing frequency

Model 3

Figure 2